

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



B36

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|----|--|
| (51) International Patent Classification 5 : C12N 15/13, 15/62 | A1 | (11) International Publication Number: WO 91/04329 (43) International Publication Date: 4 April 1991 (04.04.91) |
|---|----|--|

| | |
|---|--|
| (21) International Application Number: PCT/US90/04301 (22) International Filing Date: 1 August 1990 (01.08.90) (30) Priority data: 409,889 20 September 1989 (20.09.89) US (71) Applicant: ABBOTT LABORATORIES [US/US]; One Abbott Park Road, Abbott Park, IL 60064-3500 (US). (72) Inventor: GILLIES, Stephen, D. ; 245 Leavitt Street, Hingham, MA 02043 (US). (74) Agent: PITCHER, Edmund, R.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US). | (81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> |
|---|--|

(54) Title: METHOD OF PRODUCING FUSION PROTEINS

(57) Abstract

Disclosed are methods of producing fusion proteins including those with dual biological activities. These methods include the provision of first and second DNA sequence encoding a first and second polypeptide, respectively, the digestion of the first DNA sequence at a restriction site adjacent its 3' or 5' terminus, and the ligation of a linker/adapter sequence (1/a) to the restricted end of the first DNA sequence, thereby forming a cassette. The 1/a includes, at one end, that portion of the first DNA sequence extending from its terminus nearest the restriction site to the restriction site, and at the other end, one side of a splice site. A eucaryotic host cell is transfected with the cassette and the second DNA sequence having, at one end, one side of a splice site compatible with the side of the splice site on the 1/a. The transfected host cell is cultured to express the transfected DNA as a single chain fusion protein.

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|----|--------------------------|----|--|----|--------------------------|
| AT | Austria | ES | Spain | MC | Monaco |
| AU | Australia | FI | Finland | MG | Madagascar |
| BB | Barbados | FR | France | ML | Mali |
| BE | Belgium | GA | Gabon | MR | Mauritania |
| BF | Burkina Fasso | GB | United Kingdom | MW | Malawi |
| BG | Bulgaria | GR | Greece | NL | Netherlands |
| BJ | Benin | HU | Hungary | NO | Norway |
| BR | Brazil | IT | Italy | PL | Poland |
| CA | Canada | JP | Japan | RO | Romania |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan |
| CG | Congo | KR | Republic of Korea | SE | Sweden |
| CH | Switzerland | LI | Liechtenstein | SN | Senegal |
| CM | Cameroon | LK | Sri Lanka | SU | Soviet Union |
| DE | Germany | LU | Luxembourg | TD | Chad |
| DK | Denmark | | | TC | Togo |
| | | | | US | United States of America |

5

METHOD OF PRODUCING FUSION PROTEINS

Background of the Invention

10 This invention relates to the production of recombinant polypeptides, and in particular, to the production of synthetic proteins having plural fused domains. More specifically, this invention relates to methods of producing recombinant fusion proteins
15 having dual biological activities such as binding molecules and receptor proteins with preselected specificity and activity. Such fusion proteins are useful, for example, in imaging procedures, in the diagnosis and treatment of various human cancers,
20 infectious diseases, and dysfunctions, and in the development of vaccines.

Fusion proteins having dual activities and/or functions are known and include combinations
25 of peptide hormones, enzymes, transport and receptor proteins, viral coat proteins, interleukins, lymphokines, immunoglobulins (Igs), and fragments thereof. Some of these fusion proteins are known to provide enhanced antigenicity, and therefore, are
30 useful in the production of vaccines. Others have high biological activity, and hence, are useful in treating various infectious and deficiency diseases and disorders. Still other fusion proteins are useful as diagnostic agents because of their enhanced
35 targeting abilities. (See, e.g., U.S. 4,223,270,

-2-

U.S. 4,801,536, CA 1217156A, WO 8806630A, and JP 63267278A for examples of such fusion proteins.)

Immunoglobulin molecules have been produced 5 as fusion proteins (e.g., chimeric antibodies). Their structure is particularly conducive to the formation of fused polypeptides having a first protein domain (e.g., a variable region) from a first Ig molecule from one species having a particular 10 specificity, and a second domain (e.g., a constant region) from a second Ig of a different species (and perhaps specificity). They were developed as an alternative to non-chimeric monoclonal antibodies, many of which are of non-human (e.g., murine) origin, 15 and hence may be antigenic to humans. Human monoclonal antibodies are the most desirable therapeutic agents, as their use should provoke a greatly reduced immune response in humans. However, the production of human monoclonal antibodies by cell 20 fusion methodologies is difficult, as immunized human spleen cells are not readily available, and as human hybridomas are notably unstable.

Chimeric antibodies composed of both human 25 and non-human amino acid sequences should elicit less of an immune response in humans, and therefore should have improved therapeutic value. Accordingly, hybrid antibody molecules have been proposed which consist of Ig light (L) and heavy (H) chain amino acid 30 sequences from different mammalian sources. The chimeric antibodies designed thus far comprise variable (V) regions from one mammalian source (usually murine), and constant (C) regions from human or another mammalian source (see, for example, EPO

-3-

application nos. 84302368.0 (Genentech), 85102665.8 (Research Development Corporation of Japan), and 85305604.2 (Stanford University); P.C.T. application no. PCT/GB85/00392 (Celltech Limited); Morrison et al. (1984) Proc. Natl. Acad. Sci. U.S.A. 81:5851-6855; Boulian et al. (1984) Nature 312:643-646; and Sahagan et al. (1986) J. Immunol. 137:1066-1074).

10 The production of recombinant chimeric antibodies with predetermined specificity has typically involved the use of cloned genomic DNA fragments. For example, the genomic DNA sequences encoding H and L chains can be cloned in their 15 rearranged forms (i.e., in the DNA sequence that results from recombination events during B Cell maturation). As such, these genomic sequences contain the information necessary for their expression, (i.e. the 5' untranslated sequences, 20 promoter, enhancer, protein coding region, and donor splice site). The donor splice signals at the 3' end of the V gene segments are compatible with the splice acceptor signals at the 5' end of the Ig regions of other species. That is, the splice product between 25 the two maintains the correct reading frame. For example, when a murine V and a human C_k segment are joined and transfected into the appropriate host cell type, the primary transcript is correctly spliced and results in a mature messenger RNA (mRNA) molecule 30 with an open reading frame through both the V and C regions.

There are disadvantages to the use of genomic V region fragments for the expression of

-4-

recombinant chimeric antibodies. The first involves the cloning process itself which can be quite laborious for single-copy genes, requiring the screening of many independent clones of a phage library. Furthermore, many hybridomas contain multiple rearranged V genes which represent non-productive recombinational events. The identification of the expressed V_L or V_H segment can often require extensive DNA sequence analysis as well as confirmation by cloning and sequencing the DNA copy of the expressed mRNA (complementary or cDNA).

A more direct approach is to clone the cDNAs for both the L and H chains, and to use cDNA expression vectors for their expression. In this case, cloning is simple and rapid, since Ig mRNA is very abundant in hybridoma cells, and highly efficient methods for cDNA cloning are available. However, when one uses this approach, a problem arises when the separate expression of different V and C regions from different Ig cDNAs is desired, as in the case of chimeric antibodies having, for example, murine V regions and human C regions. The Ig cDNA represents a direct copy of the mRNA which, in turn, is a fusion of V and C exons through normal, *in vivo* RNA splicing into a continuous polynucleotide sequence. Precise excision and recombination of a murine V_H with a human $C_{\gamma 1}$, for example, is not possible because appropriate restriction sites are not present at the VC junction of both sequences.

The expression of chimeric antibodies has been accomplished through the use of cloned cDNAs. This procedure may involve the mutagenesis of

-5-

sequences in both the murine V region and human C region, near the VC junction, such that a common restriction site is created for directly joining the cDNA segments (Liu et al. (1987) Gene 54:33-40).

5

It is an object of this invention to provide a method of producing fused polypeptides. Another object is to provide methods for producing fusion proteins having dual activities and/or functions such as, for example, chimeric immunoglobulin molecules having a predetermined antigen specificity. Another object is to provide a relatively simple and rapid procedure for providing human/non-human mammalian chimeric antibodies and truncated versions thereof having reduced antigenicity in the human body. Yet another object of the invention is to provide a rapid method of producing a made-to-order chimeric immunoglobulin molecules utilizing a specifically engineered V region gene that can be attached easily to another gene, and transfected and expressed in a host cell.

Summary of the Invention

The present invention provides methods of general applicability for production of fusion 5 proteins including at least a portion of an enzyme, peptide toxin, peptide hormone, lymphokine, hormone, interleukin, and/or immunoglobulin. The invention further provides for the modification of cloned cDNAs such as Ig cDNAs so that a polypeptide such as a V 10 region with predetermined selectivity can be expressed as an independent unit or cassette fused with any one of a variety of polypeptides such as an Ig C region from the same or a different species.

15 One new method involves the construction of a linker/adapter (l/a) sequence which enables the fusion of a DNA sequence encoding a first polypeptide having a naturally occurring restriction site at or near its 3' terminus, to a DNA sequence encoding a 20 second polypeptide. The l/a sequence includes, in sequence from its 5' end, a sequence of DNA encoding the first polypeptide from a (near-)3' terminal restriction site to its 3' terminus, and a splice donor site compatible with a splice acceptor site on 25 the 5' terminus of DNA encoding a second polypeptide. In a preferred embodiment, the l/a includes a universal intron having greater than 80 nucleotide bases in length.

30 To produce the fusion protein, the DNA sequence encoding the first polypeptide is digested at a suitable restriction site near the 3' end of the gene sequence. The 5' end of the l/a is then ligated to that site, thereby producing a cassette with a

-7-

splice donor site at its 5' end. A eucaryotic host cell is transfected with the cassette as well as with a second DNA sequence encoding a second polypeptide and having the corresponding splice partner site 5 (splice acceptor) at its 5' terminus. The transfected cell is then cultured to co-express the cassette and the second DNA sequence as a single chain fusion protein.

10 Alternatively, a first DNA sequence encoding a first polypeptide may be digested at a restriction site located 3' to and adjacent its 5' terminus, thereby producing a new 5' end of that first DNA sequence. A l/a is then ligated to the new 5' end of 15 the first DNA sequence. This l/a includes, in sequence from its 3' end, DNA encoding that portion of the first polypeptide extending from a (near-)5' terminal restriction site to its 5' terminus, and a splice acceptor site compatible with a splice donor 20 site on the 3' terminus of the DNA encoding a second polypeptide. In a preferred embodiment of the invention, the l/a includes greater than about 80 nucleotide bases in length.

25 To produce the fusion protein, the DNA encoding the first polypeptide is digested at an appropriate restriction site near the 5' end of the gene. The 3' end of the l/a is then ligated to that restricted site, thereby producing a cassette with a 30 splice acceptor site at its 5' end. A eucaryotic host cell is transfected with the cassette as well as with a second DNA sequence encoding the second polypeptide and having the corresponding splice partner site (splice donor) at its 3' terminus. The

-8-

transfected cell is then cultured to co-express the cassette and the second DNA sequence as a single chain fusion protein.

5 In the case of the production of chimeric antibodies, the method includes the reconstruction of a cDNA, restricted at a naturally occurring restriction site unique to the V region-encoding portion of the cDNA, by ligation of a synthetically produced DNA sequence which replaces the removed 3' terminus of the V region and adds a 3' donor splice site that mimics the donor splice site normally found on the genomic V region-encoding DNA. This V region "cassette" can be used in conjunction with any C 10 region-encoding DNA sequence, or other protein-encoding exon, having a complementary 5' splice acceptor site. A cDNA is synthesized using reverse transcriptase and a mRNA sequence as the template which contains at least the mRNA sequence encoding 15 the Ig V region. This V region-encoding DNA sequence includes a restriction site located on the 5' side of, and adjacent, the junction (VC) of the V region sequence and the flanking sequence encoding a C region. The unique restriction site is chosen to be 20 25 a specific DNA sequence that is found rarely or only once in the V region DNA, and which is recognized by a particular restriction endonuclease capable of cleaving the DNA at that sequence.

30 Digestion at the unique restriction site leaves a fragment which encodes at least most of the V region, and preferably defines the 3' end. To the 3' end is ligated an oligodeoxynucleotide referred to herein as a l/a sequence. The l/a sequence includes

-9-

a DNA portion extending from the VC junction 5'-ward to the restriction site. In addition, the 1/a sequence includes at its 3' end a donor splice site which mimics the function of the splice site present 5 in genomic V region-encoding DNA prior to the splicing and rearranging events which occur in the nucleus during mRNA assembly. When ligated together, the restricted V region-encoding cDNA and the 1/a sequence form what is referred to herein as a V 10 region cassette.

An appropriate host cell, such as a hybridoma or myeloma, is then transfected with a vector containing the V region cassette and with DNA 15 encoding at least a portion of a polypeptide, such as a C region sequence, which includes a splice acceptor site at its 5' end and a stop signal. The splice acceptor site is compatible with the splice donor site on the 3' end of the V region cassette. The DNA 20 sequence encoding the second polypeptide may be readily retrieved from human genomic libraries. Once isolated, it may be used repeatedly to manufacture Ig having human C regions.

25 The transfected cell is then cultured to co-express the V region cassette and the C region DNA. A clone which integrates the V region cassette and the C region DNAs splices the separate V and C region DNAs to produce a mRNA encoding a synthetic 30 fused protein comprising the full length V region and a C region or portions thereof. Cotransfection with constructs encoding V_H and V_L , suitably introduced in tandem with C_H and C_L sequences, express binding proteins having an authentic Fv domain with binding

-10-

properties similar to those of the original Ig and C regions from a desired species.

In some embodiments, the resulting 5 polypeptide includes an Ig V region of murine origin and at least a portion of a human C region, thereby forming a chimeric Ig or fragment thereof. The term "chimeric Ig" is used herein to describe an Ig having amino acid sequences derived from Igs of differing 10 specificities and/or different species.

However, it should be appreciated that a given V region or portion thereof may be ligated in this fashion to polypeptides other than Ig C regions 15 to produce chimeric binding molecules other than Ig's. Furthermore, a V region need not be part of the resulting fusion protein at all.

-11-

Brief Description of the Drawing

The foregoing and other objects and features
of the invention, as well as the invention, itself,
5 may be more fully understood from the following
description when read together with the accompanying
drawings in which:

FIGS. 1A and 1B are schematic
10 representations of the methods of the invention;

FIG. 2 is the nucleic acid sequence and
deduced corresponding amino acid sequence of the
murine monoclonal 14.18 H chain V region, including a
15 leader sequence, V region, VC junction, and part of a
C region;

FIG. 3 is the nucleic acid sequence and
deduced corresponding amino acid sequence of the
20 murine monoclonal 14.18 L chain V region, including a
leader sequence, V region, VC junction, and part of a
C region;

FIG. 4 schematically depicts the
25 construction of a representative H chain V region
cassette: 4A is a schematic representation of the
murine H chain cDNA including a V and a C region; 4B
is a schematic representation of a V region cassette
including the V region cDNA sequence linked at the
30 ScaI site to an l/a DNA sequence comprising a copy of
the 3' end of the V region and a splice doner site
(ScaI-Hind III); 4C is the amino acid sequence and
corresponding cDNA sequence of a portion of the
murine H chain genomic DNA extending from the unique

-12-

restriction site (here Sca I) in the V region through the VC junction (indicated by a vertical line); 4D is the nucleic acid sequence of the l/a comprising the portion of the V region sequence extending from the 5 unique restriction site to the VC junction, and including a "universal intron" having a splice donor site adjacent the V region terminus; and 4E is a schematic representation of the l/a;

10 FIG. 5 schematically depicts the construction of a representative L chain V region cassette: 5A is a schematic representation of the murine L chain cDNA including a V and C region; 5B is a schematic representation of a V region cassette 15 including the V region cDNA sequence linked at the Afl III site to an l/a DNA sequence; 5C is the amino acid sequence and corresponding cDNA sequence of a portion of the murine L chain extending from the unique restriction site (here Afl III) in the V 20 region through the VC junction (indicated by a vertical line); 5D is the nucleic acid sequence of the l/a comprising the portion of the V region DNA sequence extending from the unique restriction site to the VC junction, and including a "universal 25 intron" having a splice doner site adjacent the V region terminus; and 5E is a schematic representation of the l/a;

FIG. 6 describes an exemplary cloning 30 strategy for the construction of a chimeric Ig expression vector: 6A is a schematic representation of the integration of the VH region cassette (consisting of a VH cassette and human C region DNA) into a pDEM expression vector; 6B is a schematic

-13-

representation of the complete H chain pDEM-VC expression vector including the V_H region cassette and human C region gene; 6C is a schematic representation of a complete L chain pDEKp-VC 5 expression vector including the V_L region cassette and human C region gene, and demonstrates the location of insertion of the V_L region cassette-human constant region fused DNA sequence into the pDEM-VC vector.

10

Description of the Invention

The method of the invention involves the preparation of a polypeptide-encoding cassette 5 including a DNA sequence which enables the splicing of that polypeptide-encoding segment to a second polypeptide-encoding DNA segment having a compatible splice sequence, thereby allowing subsequent transcription and translation of a fusion protein. 10 FIG. 1 schematically describes the methods of the present invention.

In FIG. 1A, a cassette is prepared including the reconstruction of the 3' end of a splice donor 15 (s.d.) site, and its attachment to the 3' end of a DNA sequence or exon encoding a first polypeptide (e.g., toxin, enzyme, lymphokine, interleukin, hormone, growth factor, or Ig domain gene such as a variable (V) region). The cassette is transfected 20 with expressable DNA (structural gene) for a second polypeptide (e.g., growth factor, toxin, enzyme, lymphokine, interleukin, or Ig domain such as a constant (C) region gene) having a compatible splice acceptor (s.a.) site at its 5' end. During the 25 sequence of events leading up to expression in the transfected cell, the two exons are spliced to produce a mature mRNA having a 5' end encoding the first polypeptide, and a 3' end encoding another protein domain, e.g., all or part of a human C 30 region. The resulting single chain polypeptide is a fusion of the first and second polypeptides.

In FIG. 1B, a cassette is prepared including the reconstruction of the 5' end of a splice acceptor

sit (s.a.), and its attachment to the 5' end of a DNA sequenc or xon encoding a first polypeptide (e.g., toxin, enzyme, lymphokine, interleukin, growth factor, or Ig domain such as a constant (C) region).

5 The cassette is cotransfected with expressable DNA (structural gene e.g., a growth factor, toxin, enzyme, lymphokine, interleukin, or Ig domain such as a constant region gene) having a compatible splice acceptor site at its 5' end . During the sequence of 10 events leading up to expression in the transfected cell, the two exons are spliced to produce a mature mRNA having a 3' end encoding the first polypeptide, and a 5' end encoding another protein domain. The resulting single chain polypeptide is a fusion of the 15 first and second polypeptides.

For example, an Ig V region cassette may be spliced to a C region immunoglobulin sequence, resulting in the expression of an intact Ig molecule, 20 an H or L chain, or fragment thereof. Alternatively, a toxin-encoding cassette may be spliced to an Ig V region, resulting in a "magic bullet"-type therapeutic agent.

25 During the sequence of events leading up to expression in the transfected cell of the fusion protein, mRNA derived from the two DNA sequences are spliced to produce a mature mRNA having a 5' end encoding a complete V_H or V_L and a 3' end encoding 30 another protein domain, e.g., all or part of a human C region. The resulting single chain polypeptide is a fusion of the V region and the polypeptide encoded by the exon of the structural gene. For example, a V region cassette may be spliced to a C region

-16-

sequence, resulting in the expression of an Ig molecule, an H or L chain, or fragment thereof.

A representative case includes the
5 construction of a chimeric binding protein such as an antibody. According to the method described in FIG.
1A, an Ig V region cassette is prepared including a reconstruction of the 3' end of the V region and attachment of a splice donor site to the 3' end of a
10 V_H- and/or V_L-encoding cDNA sequence (first polypeptide). The modified V region cDNA is cotransfected with a structural gene for an Ig constant region (second polypeptide) having a splice acceptor site at its 5' end.

15

The cDNA encodes a H or L chain V region (most likely non-human, e.g., murine) of a defined specificity, while the C region exon(s) encode the H or L chain C region of another (most likely human) Ig species. Expression of H and L constructs in a single competent host cell results in production of intact chimeric immunoglobulins having a desired specificity and, for example, an intact human constant region. Useful V regions include the V_H and
25 V_L domains of murine monoclonal antibody 14.18 (Mujoo et al. (1987) Cancer Res. 47: 1098-1104) which recognizes the disialoganglioside GD₂ on the surface of many neuroblastoma, melanoma, glioma, and small lung carcinoma lines and tissues. This V region can
30 be combined with various C regions such as that of the Ig human gamma (H) and kappa (L) chains. Alternatively, the V region may be of human origin.

To produce the V region-containing fused polypeptide, a V region cassette is constructed and placed in an appropriate vector, together with, for example, a C region-encoding DNA sequence.

5

The V region cassette is constructed as follows. cDNA is synthesized from mRNA isolated from cells rich in V region-encoding mRNA. Useful cell sources include lymphoid cells such as lymphocytes, 10 myelomas and hybridomas. A number of established protocols are available for mRNA isolation and cDNA synthesis (see, e.g., Maniatis et al., *ibid.*). The longest cDNA molecules are identified, for example, with specific C_H (C_{Y3}), C_L (C_K) sequences, or other 15 nick-translated or oligonucleotide probes in order to detect either or both the L chain or H chain C region. They are then sequenced to identify the DNA sequence of the 3' end of the V region and to locate unique restriction sites within the V region as near 20 as possible to the VC junction.

The next step is to synthesize an oligonucleotide which is used to recreate the sequence between the unique restriction site and 25 splice donor site. In the intact cDNA, the DNA sequence just 5' of the VC junction constitutes the 5' half of the native splice donor site from the genome of the source cell, while the sequence just 3' of the junction (in the C region) represents the 3' 30 half of the native splice acceptor site. The missing portion of each native splice site cannot be deduced from the cDNA since they were removed as part of the introns during mRNA synthesis in the source cell nucleus during the splicing events.

In accordance with the invention, a DNA fragment is synthesized which restores the 3' side of the donor splice site, or alternatively, the 5' side 5 of the acceptor splice site. The fragment should include enough DNA (at least about 80 nucleotide bases) to insure that efficient splicing will occur with an appropriate splice partner sequence. In addition, it is essential that a splice donor/acceptor 10 pair be used that results in a splice maintaining the original reading frame.

To this end, a 1/a DNA sequence has been synthesized which includes an intron having parts of 15 its sequence in common with many introns including a splice donor sequence at its 3' end. This universal intron (U) (see FIG. 1) includes the splice donor sequence GTAATGTG; however, other sequences are equally as useful as splice donors and could be 20 utilized as well. A partial list of exemplary useful splice donor sequences (5' and 3' portions) is shown in TABLE 1. These sequences are from known murine V_H and V_L genomic DNAs. The 5' ends are found at the VC junction of various murine cDNAs, while the 3' ends 25 are removed during the splicing processes. The sequence of the universal intron used herein is also listed for comparison. While the sequences immediately preceding the VC junction should not differ from those shown (since mutations here might 30 lead to RNA splicing problems), it should be mentioned that a 3' portion of the V region in many expressed Ig genes (namely, the J or joining region) undergo somatic mutation. These changes should also

-19-

be included in the synthetic l/a sequence, since this region may contribute to antigen binding.

TABLE 1

5

Murine Ig V(J) Region Splice Donor Sequences

| | <u>5' end</u> | <u>3' end</u> |
|-----------------------|--------------------|---------------|
| 10 J _{k1} |GCTGGAAATCAAAC | GTAAGTAG.... |
| J _{k2} |GCTGGAAATAAAAC | GTAAGTAG.... |
| J _{k3} |GTTGGAAATAAAAC | GTAAGTAG.... |
| J _{k4} |GCTGGAGCTGAAAC | GTAAGTAC.... |
| 15 J _{H1} |CACCGTTTCCTCAG | GTAAGCTG.... |
| J _{H2} |CACAGTCTCCTCAG | GTGAGTCC.... |
| J _{H3} |CACTGTCTCTGCAG | GTGAGTCC.... |
| J _{H4} |CACCGTCTCCTCAG | GTAAGAAT.... |
| 20 (Universal Intron) | | GTAATGTG.... |

The same universal intron sequences have been used successfully with other V regions. In all 25 cases the Ig L and H chains were efficiently expressed, demonstrating the generality of this process.

Recombinant DNAs comprising a truncated cDNA 30 with restored 3' terminal fragment and donor splice site encoding a first polypeptide and are expressed with another DNA having a splice acceptor site 5' of a region encoding a second polypeptide. For example, a splice acceptor may be derived from genomic Ig DNA

-20-

encoding all or a portion of an immunoglobulin C region. If a 1/a is used which includes a universal intron having a splice acceptor site, any such known splice acceptor site may be incorporated into the 5 sequence. Typically, the 5' end of the splice acceptor site includes a pyrimidine-rich region (i.e., T, C), followed by the sequence AG and the splice junction. Therefore, the universal intron may include such a splice acceptor sequence at its 5' end 10 if the 1/a to be used is to be linked to the 5' end of the first DNA sequence instead of the 3' end. Exemplary sequences comprising suitable splice acceptor sites are listed in TABLE 2. These sequences are the acceptor sequences of the: (1) CH₁, 15 (2) hinge, (3) CH₂, and (4) CH₃ exons of the C_γ1 gene. However, many other suitable splice acceptor sequences can be found in the literature.

TABLE 2

20

Murine Ig Splice Acceptor Sites

| | <u>5' end</u> | <u>3' end</u> |
|----|---------------|---------------|
| | (1) CTCTTGCAG | CCTCC |
| 25 | (2) TCTCTGCAG | AGCCC |
| | (3) CTTCCTCAG | CACCT |
| | (4) GTCCTACAG | GGCAG |

Upon co-expression of the two DNAs, the host 30 cell nuclear enzymes produce mRNA by implementing the normal splicing events, resulting in an mRNA encoding a fused protein. In the case of a chimeric binding protein, the mRNA may encode (5' to 3') a V_H or V_L domain, which may be identical to the native sequence

-21-

up to the VC junction, attached directly to another polypeptide such as at least a portion of at least one C region domain, which for example, may comprise human sequences. The 3' half of the donor splice 5 site (and any nucleotides downstream) and the 5' half of the acceptor splice site (and any nucleotides upstream) are removed as an intron, resulting in an mRNA encoding the properly fused protein.

10 Other suitable sources for C-region encoding DNA having an attached 5' splice acceptor site include human or other mammalian genomic libraries. Alternatively, a splice acceptor site having a suitable sequence may be ligated to the 5' 15 end of any desired expressible DNA sequence which may then be fused to a DNA sequence encoding a V domain having a specificity limited only by the specificity of available monoclonal antibodies or those producible by known techniques.

20 Normally, in the case of chimeric Ig molecules, both exons are placed on the same vector under control of a single regulating sequence. Also, it is preferred to coexpress both L and H chain 25 constructs so that the host cell secretes an intact fusion protein. The method requires use of a host cell having the enzymes which recognize the DNA splice signals and effect proper splicing.

30 Particular vector construction, host cell selection, transformation, and methods of expression do not, per se, constitute an aspect of the invention, but can be selected and implemented by skilled workers based on personal preference and

convenience. Techniques adaptable for use in the invention are disclosed, for example, in Current Protocols in Molecular Biology (Greene Publishing Associates, 430 Fourth Street, Brooklyn, N.Y., 5 1989). Useful vectors include any number of known plasmids which contain the correct signals for transcription and translation of the genes of interest. Enhancer elements may be present, and additional signals for polyadenylation and splicing 10 must be present in cases where they are not provided by the gene itself. For example, all of the signals for the expression of functionally rearranged Ig genes are present on a continuous stretch of DNA and include the transcription promoter, the 15 polyadenylation and termination sites, and the splicing signals for excision of the intron sequences. Additional information that must be provided by the vector is a selectable marker gene. This gene must also contain the signals for 20 expression of the selectable phenotype (usually resistance to the lethal effect of a toxic drug such as methotrexate, for example). Therefore, if the vector encodes both the first and second polypeptides, it is necessary that it provide the 25 sequence information for three separate transcription units in a limited amount of space.

The recombinant cassette-containing vector is transfected into an appropriate host cell. The 30 choice of host cell line, in addition to the criterion noted above, is based on its ability to grow in a growth media, preferably one that is commercially available and serum-free media as well as its ease of selectivity after transformation. For

the production of chimeric antibodies, useful host cells include myelomas or hybridomas such as, for example, the murine non-Ig-producing Sp2/0 Ag 14 hybridoma cell line. Useful cells are widely 5 available in repositories and from commercial sources and may be isolated readily from natural sources by those skilled in the art.

One method for introducing recombinant DNA 10 into cells is electroporation (see, e.g., Potter et al. (1984) Proc. Natl. Acad. Sci. USA 81:7161-7165), which requires specialized equipment and the availability of highly purified DNA. However, many 15 different lines can be transformed using this method if conditions are optimized for the specific cell type.

Another transfection method is protoplast (spheroplast) fusion (see, e.g., Sandri-Goldin et al. 20 (1981) Molec. Cell. Biol. 1:743-752). Bacteria harboring the recombinant plasmid of interest are fused to the lymphoid cells with a chemical agent, generally 45-50% polyethylene glycol in a buffered, isotonic solution. This method is simple and does 25 not require extensive purification of plasmid DNA. In addition, very high transformation frequencies can be obtained, and the time for obtaining highly productive transfected cell clones is reduced because this transfection method is likely to give 30 transfectants containing multiple copies.

Cells which are successfully transformed with the cassette-containing vector must then be isolated from those which are not. Many methods are

available for the selection of transfected cells. For example, the guanine phosphoribosyl transferase (gpt) and neomycin resistance markers may be used for selection purposes in lymphoid cells. The gene 5 encoding the marker would be included on the V-region encoding vector. The resistant form of dihydrofolate reductase (DHFR) can also be used for the selection of hybridoma cell transformants as well as for subsequent amplification of the marker and flanking 10 product genes.

The transfected cell is then cultured to express the polypeptide encoded by the cassette. Culturing may be in vitro, or in the case of 15 recombinant antibodies, may be accomplished by employing other strategies such as in vivo culturing in ascites fluid.

Improvements in the productivity of 20 transfected cells is possible through the use of multiple subcloning steps in the presence of MTX when the sequences of interest are co-expressed with the DHFR marker. For example, after two cycles of subcloning by limiting dilution, cells are obtained 25 that are capable of producing antibody, in spent suspension culture medium, at levels of 35-100 µg/mL. These levels of expression are greatly reduced if the cells are passaged in the absence of MTX. It is possible, however, to maintain stock 30 cultures in the presence of the drug and then omit it from the last scale-up step. The final yields of antibody in this case are not diminished by the omission of MTX.

-25-

In addition to open suspension culture methodologies, other scale-up perfusion technologies such as hollow fiber reactors (see, e.g., Von Wedel (1987) in Commercial Production of Monoclonal Antibodies)

5 Antibodies: A Guide for Scale Up (Seaver, ed.) Marcel Dekker, Inc., New York) and microencapsulation (see, e.g., Rupp (1986) "Use of Cellular Microencapsulation in Large-Scale Production of Monoclonal Antibodies" in Large Scale Mammalian Cell Culture (Tolbert and 10 Feder, eds.) Academic Press, New York) are especially useful with the present expression system. For example, it is possible to express high levels of a recombinant human antibody within microcapsules using murine hybridoma transfectants. The entrapped cell 15 cultures can be maintained in medium containing greatly reduced levels of fetal bovine serum, and in some cases, can be maintained in completely serum-free media.

20 Perfusion methods not only allow for higher cell densities (which may promote cell-cell interactions and thereby reduce the serum requirement), but may also be useful for the removal 25 of MTX from the culture prior to purification of the antibody. This is possible using microencapsulation culture wherein the semipermeable capsule membrane is made to retain the high molecular weight antibody molecule but allow for the diffusion out of the capsule of smaller molecules. At the end of a 30 perfused culture run, long after cell division has stopped but production of antibody continues, the culture could be maintained without MTX. Further removal of the drug is possible by washing and

dialyzing the capsules in physiological saline prior to their disruption.

The invention will be further understood
from the following non-limiting examples.

EXAMPLE

Two separate H and L chain cDNA libraries
10 are prepared according to the method of Gubler et al.
(Gene (1983) 25:263-269), herein incorporated as
reference. Double-stranded, blunt-ended cDNA is
synthesized from polyadenylated polyA mRNA isolated
from the murine hybridoma cell line, 14.18.

15

A double-stranded polylinker with the sequence:

5' HO-AATTCCCTCGAGTCGAC
GGAGCTCAGCTG-PO₄ 5'

is then ligated to the cDNA. This linker serves several purposes. The first allows the blunt-ended cDNA to be cloned into the EcoRI site of λgt10 phage DNA via the AATT sticky-end sequence. Note that only the 5' blunt end has been phosphorylated for ligation to the blunted cDNA. By not phosphorylating the 5' EcoRI sticky end, polymerization of the linker cDNA does not occur and subsequent enzymatic digestion after linker ligation is unnecessary. This, in turn, makes it unnecessary to methylate internal EcoRI sites which would otherwise be cut in the process. The 5' phosphorylated sticky end provided by the EcoRI digested λgt10 DNA is sufficient for ligation and subsequent cloning of the cDNA in the phage host.

Another function of this linker sequence is to provide XhoI (CTCGAG) and SalI (GTCGAC) sites for subsequent manipulations. These sites rarely occur 5 in murine Ig cDNAs and thus provide unique restriction sites at the 5' and 3' ends. Since both sites provide the same 5' overhang sequence upon restriction (TCGA), either can be used in expression vectors with XhoI cloning sites. If one site happens 10 to appear in the cDNA sequence, the chances are small that the second site would also be present.

The linkered cDNA is then fractionated by polyacrylamide gel electrophoresis (PAGE) to enrich 15 for full-length L and H chain cDNAs. Two fractions of cDNA corresponding to full-length L chain (900-1100 bp) and H chain (1400-1600 bp) can be separately isolated. The DNA in each gel fraction is separately eluted and ligated to EcoRI-digested 20 λgt10 DNA. Following in vitro packaging with commercially available packaging mix (Stratagene, San Diego, CA), recombinant phage is plated and screened by filter hybridization using various C region probes.

25 Ten phage clones from each screening are analyzed further by restriction analysis using EcoRI. This is most rapidly achieved by preparing small-scale phage lysates which provide enough DNA for restriction digest and determination of the 30 length of the cDNA insert. Clones that appear to be full-length are labelled with radioactivity, e.g. at the unique EcoRI sites at the 5' and 3' ends (relative to the original mRNA polarity). Following digestion with a second enzyme, the clones are

sequenced by the method of Maxam and Gilbert (Meth. Enzymol. (1980) 65:499-559), herein incorporated by reference, and screened with probes for murine C_{y3} and C_x sequences.

5

The longest H cDNA sequence is shown in FIG. 2. The H chain clone appears to be very close to full-length if it is assumed that the second ATG in the sequence represents the true initiation codon. Although both ATG codons are in the correct reading frame, the first is probably too close to the 5' end of the mRNA for efficient initiation and, if translated, would encode an uncharacteristically long leader sequence. The use of the second ATG codon would result in the synthesis of a very typical Ig leader sequence of 19 amino acids. Because additional 5' untranslated sequences are to be added to this cDNA in the expression vector (see below), the first ATG would no longer be at the 5' end of the resulting fusion mRNA, thereby increasing the likelihood of the translation of the aberrant leader sequence. To avoid this problem, the cDNA sequence is truncated by limited Bal31 exonuclease treatment. To the resulting modified cDNA clone (H2c) is then attached an XhoI linker (see FIG. 2). The expression of this cDNA should result in an mRNA encoding a normal Ig leader sequence. The remainder of the V_H portion of the cloned cDNA appears to be a normal, functional variable region.

30

Examination of the L chain clone (FIG. 3) shows a typical L chain sequence encoding a 19 amino acid leader followed by sequences that are highly homologous to the anti-GAT family of V_K genes. The

murin κ C region begins with th arginine residue at position 114.

Reconstruction of the donor splice sites on 5 the 3' ends of the cloned cDNAs for V_H and for V_L is carried out as shown in FIGs. 4 and 5. The first step is to sequence both V_H and V_L regions, and then to identify unique restriction sites near the junctions of the V and C regions (see FIGs. 2 and 3).

10

The sequence between the unique restriction site and the original splice donor site in the genomic V gene segment is then recreated using the universal intron sequence previously shown in TABLE 15 1. For convenience, the oligonucleotides used to reconstruct the intron sequence are cloned as a SnaB1 to HindIII fragment (the latter site was not originally part of the intron). The SnaB1 site can then be cut to produce a blunt end beginning with GTA 20 (the bases that reconstitute the splice signal) while the HindIII site serves as the 3' cloning site for insertion of the cassette into the expression vector.

The small linker portions of the V regions 25 shown in FIGs. 4 and 5 are synthesized by the hybridization and ligation of overlapping oligonucleotides to produce fragments with the appropriate "sticky" 5' end, as well as a blunt 3' end. Each synthetic 1/a sequence is then cloned into 30 a plasmid vector for sequence verification. In some cases the 5' cloning site may not be unique or present in a convenient plasmid. In these cases, additional restriction sites can be added to the 5' end of the synthetic DNA for cloning purposes. After

DNA sequencing, the 1/α fragments are ligated to the remaining portions of the V regions. The complete V_H and V_L region cassettes are then cloned as XbaI to HindIII fragments.

5

The mammalian cell expression vector used for the expression of the chimeric 14.18 antibody is shown in FIG. 6. It was derived from the pDEM vector (FIG. 6A), which has been used to express cDNA sequences in lymphoid cells. This vector was designed as follows. At the left end of the schematic linear representation there is the selectable marker gene, DHFR, which is oriented so that transcription would proceed from right to left. 10 This transcription unit is composed of SV40 regulatory signals (enhancer and promoter), a cDNA encoding the resistant form of mouse DHFR, and an SV40 polyA site. 15

20 Adjacent to this gene (to the right of the SalI site) is the transcription unit for the cDNA of interest. The first block of sequence represents the Ig H chain enhancer (E_H), the sequence that is normally involved in the active transcription of Ig 25 H-chain genes. The enhancer sequence is followed by the promoter for the murine metallothionein (MT) gene. Many other promoters can be used in this position including those from Ig genes, as in the L chain transcription unit shown in FIG. 6C. A 30 synthetic restriction site (for XbaI) is placed in the 5' untranslated region of the MT sequence so that cDNAs can be inserted at this position and be transcribed as fusion mRNAs from the MT promoter. In this case it is necessary that the cDNA contain its

-31-

own translation initiation codon since it is not provided by the vector. Beyond the XhoI site (after the insertion site for cDNAs) is a segment derived from the 3' untranslated region of the murine C κ 5 region that contains the polyA addition site. A cDNA expressed from this transcription unit would then be composed of the 5' untranslated sequence from the MT mRNA, the specific coding sequence of interest, and the 3' untranslated sequence normally found on murine 10 κ mRNAs.

In the present example of chimeric antibodies, some of the components of this plasmid are superfluous, namely the 3' untranslated sequence 15 and the polyA site from the κ gene segment. Instead of these components, the V region cassette and the human C region gene segment between the XhoI and SalI sites of plasmid pDEM (see FIG. 6A) are inserted. Since the C region gene already provides the 3' 20 untranslated sequences and the polyA site, the only signals to be used from the plasmid are those necessary for the transcription of the cDNA (enhancer and promoter). The resulting plasmid is shown in FIG. 6B.

25

A similar construct is made for the L chain, but in this case the plasmid has a guanine phosphoribosyltransferase (gpt) selectable marker gene not present in the final vector, and the C 30 region is that of the human κ gene. The promoter used for L chain cDNA expression is derived from a V κ gene, and a unique XhoI site is placed in the 5' untranslated region for fusion to the cDNA. This latter construct is designed so that the entire L

chain transcription unit is flanked by SalI sites, while in the case of the H chain vector, only a single SalI site is retained. In this way it is possible to excise the entire L chain transcription unit (about 3 kb), and then to insert it into the unique SalI site of the H-chain expression vector (FIG. 5C).

Another important design feature of the Ig 10 expression vector involves the use of the E_{μ} enhancer for the expression of both H and L chains. Since this enhancer is much more powerful than the L chain counterpart (at least in transfection experiments), a higher and more balanced level of 15 expression of both chains is possible. The L chain transcription unit also contains some additional sequences between the SalI site and the E_{μ} sequence. This sequence is derived from the λ_1 L chain promoter, and serves to attenuate the 20 bi-directional enhancing effect of E_{μ} in one direction (away from the L chain transcription unit) without affecting enhancement of transcription of the cDNA (EPO 87/300658.9). When the L chain transcription unit is inserted into the expression 25 vector as shown in FIG. 6, this sequence prevents the overexpression of the adjacent marker gene by blocking the enhancer effect in that direction. In this way it is hoped that transfectedants will produce more of the transfected protein of interest relative 30 to the product of the marker gene.

The murine non-Ig-producing hybridoma cell line, Sp2/0 Ag 14 cells are transfected with the chimeric Ig plasmid construct described above using a

-33-

modified protoplast fusion method (Gillies et al. (1983) Cell 33:717-728), herein incorporated by reference.

5 After plating in 96-well culture dishes, selection medium (Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum) containing MTX (0.1 μ M) is added. Cells are fed with 50% culture medium replacement at 3-4 day intervals for a total
10 of three feedings. Colonies of MTX-resistant cells appear in 10-14 days.

At this point culture supernatants are assayed for secreted human antibody determinants by
15 an ELISA assay. Nearly all the MTX-resistant colonies secrete significant levels of human antibody into the medium.

After identifying the clones which secrete
20 the most antibody, the expression level is enhanced by increasing the concentration of MTX in the medium. The cells quickly adapt to dramatic increases in concentration (as much as 10-fold in a single step) with little or no cell death. The level
25 of MTX is increased from 0.1 to 10 μ M in a period of about three weeks. During this time, the concentration of antibody increases from a range of 2 to 8 μ g/mL (at 0.1 μ M
MTX) to a range of 10 to 40 μ g/mL of spent culture
30 medium (at 10 μ M MTX). Levels in excess of 100 μ g/mL of spent culture medium have been obtained using this and other similar constructs embodying the invention.

-34-

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The present embodiments are therefore to be considered in all 5 respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are 10 therefore intended to be embraced therein.

We claim:

1. A method of producing a fusion protein comprising the steps of:

5 (a) providing a first DNA sequence encoding a first polypeptide and defining a restriction site located 5' to and adjacent the 3' terminus thereof;

10 (b) digesting said first DNA sequence at said restriction site to produce a 3' end;

15 (c) ligating a linker/adapter (l/a) DNA sequence to said 3' end to produce a cassette, said l/a DNA sequence comprising, in sequence from its 5' end:

20 DNA encoding a portion of said first polypeptide extending from said restriction site to said 3' terminus of said first DNA sequence; and

a splice donor site;

25 (d) transfecting a eucaryotic host cell with said cassette and with a second DNA sequence, said second DNA sequence comprising:

30 DNA encoding a second polypeptide; and

5 DNA defining a splice acceptor site at
the 5' terminus of said second DNA
sequence, said splice acceptor site
being compatible with said splice donor
site on said 1/a DNA sequence; and

(e) culturing said transfected cell to
co-express said cassette and said second DNA
sequence as a single chain fusion protein.

10

2. A method of producing a fusion protein
comprising the steps of:

15

(a) providing a first DNA sequence encoding
a first polypeptide and defining a
restriction site located ~~in~~ to and adjacent
the 5' ~~terminus~~ thereof;

20

(b) digesting said first DNA sequence at
said restriction site to produce a 5' end;

25

(c) ligating a linker/adapter (1/a) DNA
sequence to said 5' end to produce a
cassette, said 1/a DNA sequence comprising,
in sequence from its 3' end:

30

DNA encoding a portion of said first
polypeptide extending from said
restriction site to said 5' terminus of
said first DNA sequence; and

a splice acceptor site;

(d) transfected a eucaryotic host cell with said cassette and with a second DNA sequence, said second DNA sequence comprising:

5

DNA encoding a second polypeptide; and

10 DNA defining a splice acceptor site at its 3' terminus, said site being compatible with said splice donor site on said 1/a DNA sequence; and

15 (e) culturing said transfected cell to co-express said cassette and said second DNA sequence as a single chain fusion protein.

3. The method of claim 1 or 2 wherein said providing step (a) comprises providing a complementary DNA (cDNA) sequence encoding said first 20 polypeptide.

4. The method of claim 1 wherein said 1/a DNA sequence ligated in step c) comprises greater than about 80 nucleotide bases.

25

5. The method of claim 1 wherein said 1/a DNA sequence ligated in step c) comprises a universal intron including greater than about 80 nucleotide bases.

30

6. The method of claim 1 or 2 wherein said first polypeptide is selected from the group consisting of growth factors, enzymes, hormones, lymphokines, interleukins, immunoglobulins, analogs 5 thereof, and fragments thereof.

7. The method of claim 1 or 2 wherein said fusion protein comprises a chimeric fusion protein.

10 8. The method of claim 1 wherein said first DNA sequence provided in step a) comprises a DNA sequence encoding a non-human immunoglobulin V region, and said fusion protein comprises a chimeric binding protein.

15

9. The method of claim 2 wherein said second DNA sequence provided in step c) comprises a DNA sequence encoding a non-human immunoglobulin V region, and said fusion protein comprises a chimeric 20 binding protein.

10. The method of claim 8 wherein said first DNA sequence provided in step a) comprises a DNA sequence encoding a murine immunoglobulin V region.

25

11. The method of claim 9 wherein said second DNA sequence provided in step c) comprises a DNA sequence encoding a murine immunoglobulin V region.

30

12. The method of claim 8 wherein said second polypeptide comprises at least a portion of an immunoglobulin constant (C) region, and said chimeric binding protein comprises at least a portion of a chimeric immunoglobulin.

13. The method of claim 9 wherein said first polypeptide comprises at least a portion of an immunoglobulin constant (C) region, and said chimeric binding protein comprises at least a portion of a chimeric immunoglobulin.

14. The method of claim 12 wherein said second polypeptide comprises at least a portion of a human immunoglobulin constant (C) region.

15. The method of claim 13 wherein said first polypeptide comprises at least a portion of a human immunoglobulin constant (C) region.

15

16. A method of producing a chimeric binding protein comprising the steps of:

20 (a) providing a DNA sequence encoding an immunoglobulin variable (V) region and defining a restriction site located 5' to and adjacent the junction of said V-region-encoding DNA sequence and DNA encoding an immunoglobulin constant (C) region;

25 (b) digesting said V region-encoding cDNA at said restriction site to produce a 3' end;

30 (c) ligating a linker/adapter (1/a) DNA sequence to said 3' end to produce a V region cassette, said 1/a DNA sequence comprising, in sequence from its 5' end:

a DNA encoding a portion of said V region-encoding cDNA extending 5' from said restriction site to said VC junction; and

5

a splice donor site;

10 (d) transfected a eucaryotic host cell with said cassette and with a DNA sequence comprising:

15 DNA defining a splice acceptor site at its 5' end, compatible with said splice donor site on said l/a DNA sequence and DNA encoding a polypeptide; and

20 (e) culturing said transfected cell to co-express said cassette and said polypeptide-encoding DNA sequence as a single chain, fused, chimeric binding protein.

17. The method of claim 16 wherein said DNA provided in step a) comprises complementary DNA
25 (cDNA).

18. The method of claim 16 wherein said DNA provided in step a) comprises DNA encoding an immunoglobulin light chain V region.
30

19. The method of claim 18 wherein said DNA encodes a non-human immunoglobulin domain or portion thereof.

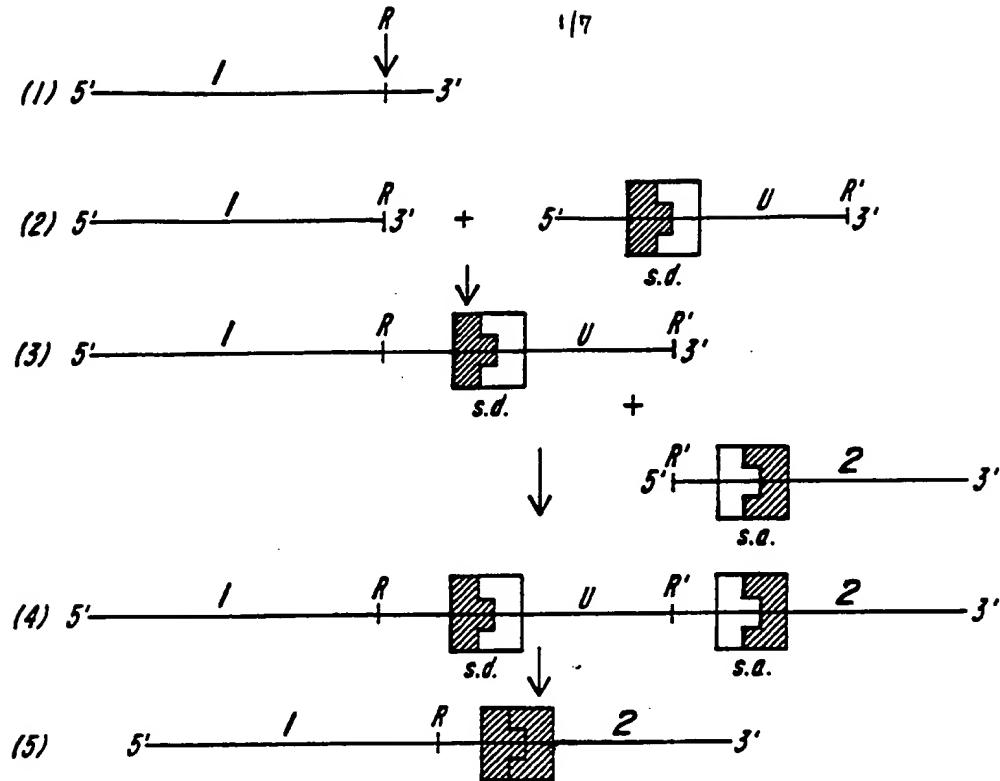
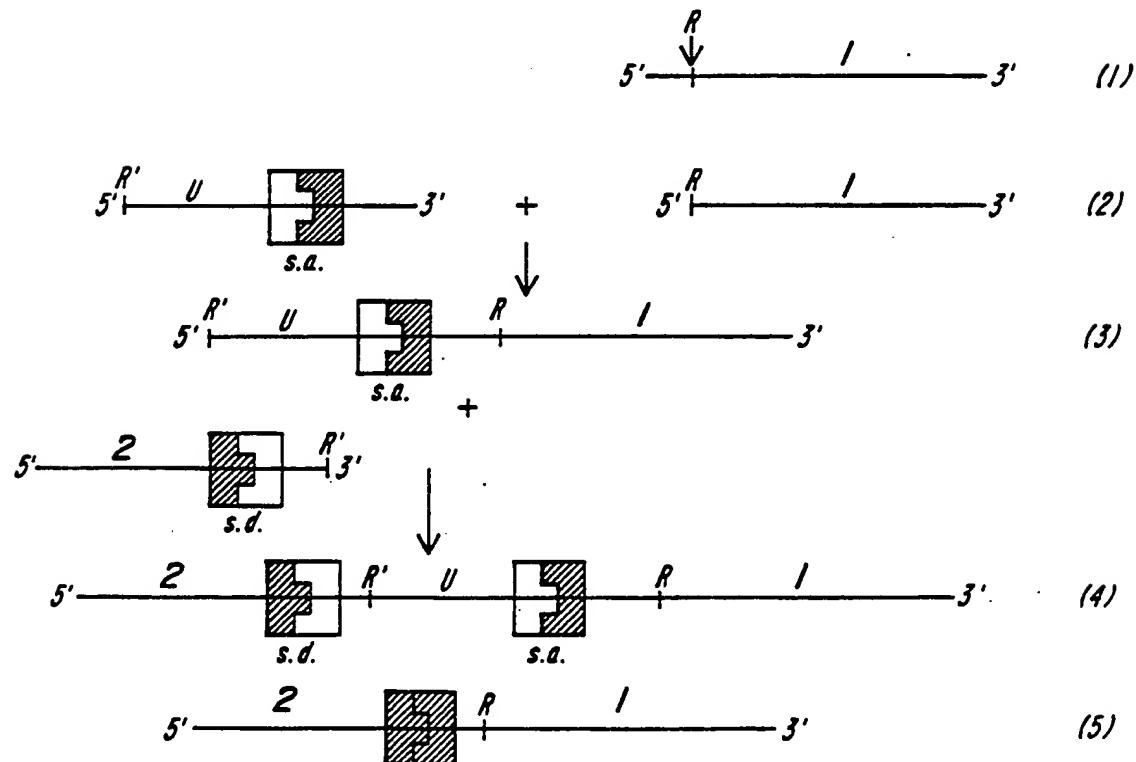
20. The method of claim 19 wherein the cDNA encodes a murin immunoglobulin domain or portion thereof.

5 21. The method of claim 16 wherein said DNA provided in step a) comprises DNA encoding an immunoglobulin heavy chain V region.

10 22. The method of claim 16 wherein said 1/a DNA sequence ligated in step c) comprising greater than about 80 nucleotide bases.

15 23. The method of claim 16 wherein said 1/a DNA sequence ligated in step c) comprising a universal intron including greater than about 80 nucleotide bases.

20 24. The method of claim 16 wherein said transfecting step d) comprises transfecting a host cell with DNA encoding a polypeptide, said polypeptide including at least a portion of a human immunoglobulin C region.

**FIG. 1A****FIG. 1B**

2 / 7

| | | | |
|---|---|---|---------------------------------|
| <pre> H2c ↓ <u>ATGCTCTCACAGTCCGAGACACTGACTAATGACTGACCTGATGATGTTATCCTGGTAACTACGGTCCTCTG MetGlyTrpThrIleLeuIleSerValThrGlyValSerGlu +1 Leader </u> </pre> | <p>100</p> <p>150</p> <p>200</p> <p>250</p> <p>300</p> <p>350</p> <p>400</p> <p>450</p> | <pre> 50 GCGACTCTGCTGACTGACCTGAGCCCTAGGGCTCAGGTCACTGATCATATCCGACCCCTGCTCATTTACTGCTACACCTG ValGlnLeuIleGlnSerGlyProGluLeuGlySerAlaSerAlaSerValMetIleSerCysLysAlaSerGlySerPheThrGlySerAlaSerAla 10 20 30 40 50 60 70 80 90 100 </pre> | <pre> H4 Constant Region </pre> |
|---|---|---|---------------------------------|

FIG. 2

3/7

Sequence alignment diagram showing the relationship between the Leader sequence and the Constant Region.

The Leader sequence is aligned with positions 1-150, and the Constant Region is aligned with positions 150-500.

The Leader sequence starts at position +1 and ends at position 150. The Constant Region starts at position 150 and ends at position 500.

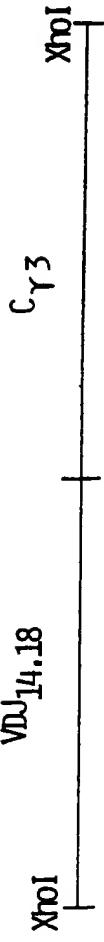
A vertical line with arrows at both ends connects the two sequences.

The Leader sequence is labeled "Leader" and has a scale from 100 to 50.

The Constant Region is labeled "Constant Region" and has a scale from 500 to 100.

The alignment shows several conserved regions and some differences, such as a proline residue in the Leader at position 100 which is replaced by a glycine in the Constant Region at position 450.

FIG. 3

FIG. 4A**FIG. 4B****FIG. 4C**

5' Met Glu Tyr Trp Gly Gln Gly Thr Ser Val Thr Ser Ser Ala
 ATG GAG TAC TGG GGT CAA GGA ACC TCA GTC ACC GTC TCC TCA G
 ScaI

3'

V_{DJ} C_γ
 ScaI HindIII

Linker/Adaptor (Splice Donor Site)

4/7

Murine cDNA ...

FIG. 4D

GAG TAC TGG GGT CAA GGA ACC TCA GTC ACC GTC TCC TCA G
 ScaI

Splice Donor

HindIII

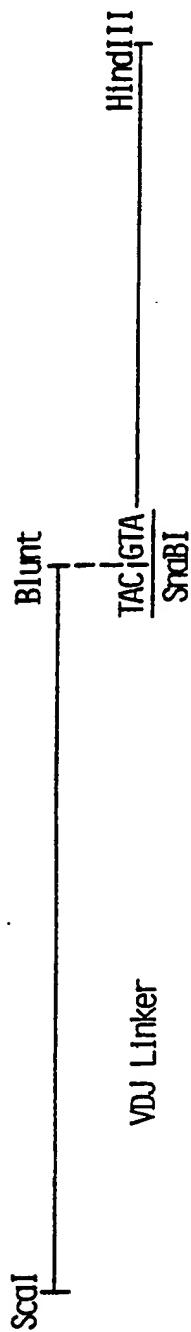
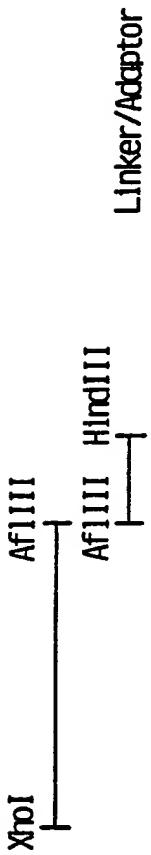
FIG. 4E

FIG. 5A

cDNA

*FIG. 5B*

5'

*FIG. 5C*A CAC GTT CCT CGG CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA CAf1111

VJ

C

3'

Thr His Val Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Arg Ala Asp Ala Ala
ACA CAT GTT CCT CGG CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA C
Af1111

cDNA

5/7

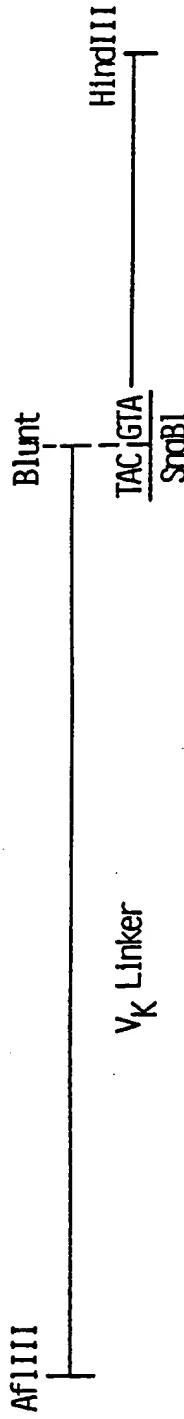
Linker/Adaptor

*FIG. 5D*A CAC GTT CCT CGG CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG AAA CAf1111

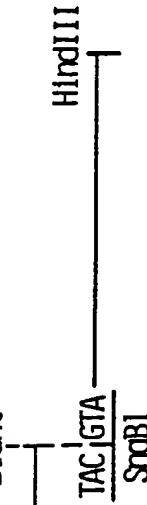
GTAAGTG //GGTAATGTCACTAAGCTT
HindIII

Splice Donor

Af1111

*FIG. 5E*

Blunt



Universal Intron

6/7

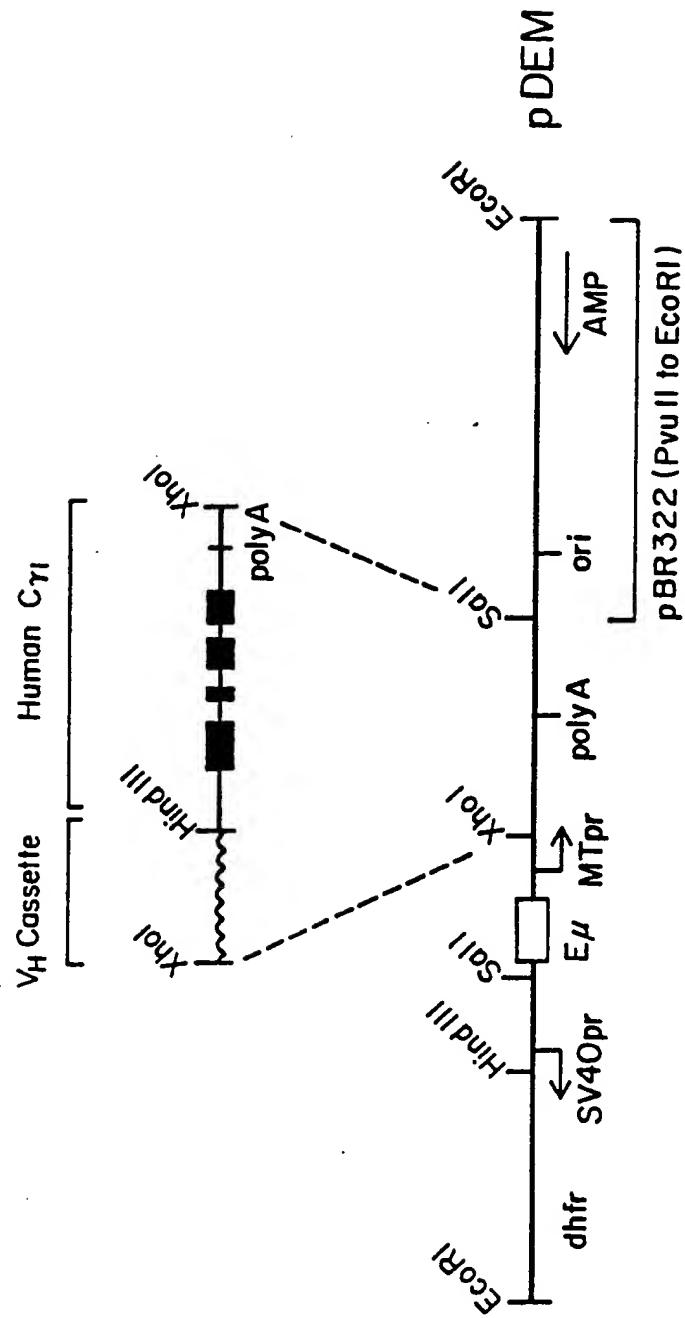


FIG. 6A

7/7

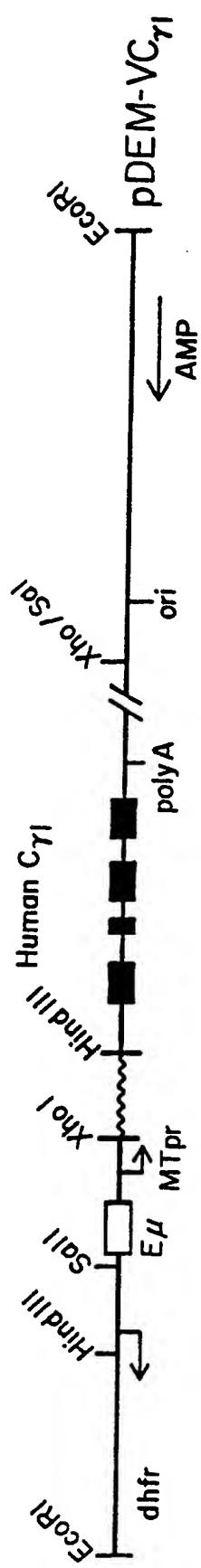


FIG. 6B

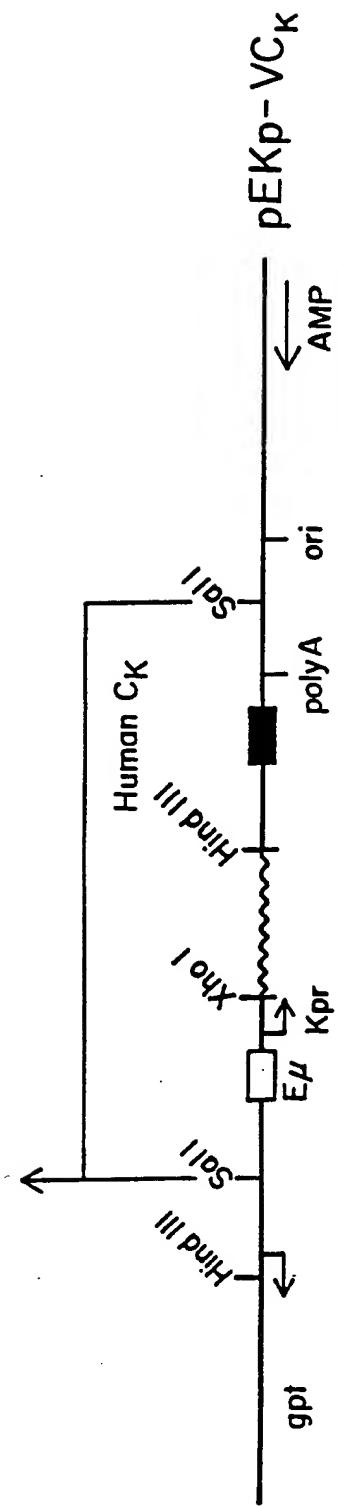


FIG. 6C

SUBSTITUTE SHEET

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/04301

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all):

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12N 15/13, 15/62

U.S.CL: 430/69.7; 435/172.3

II. FIELDS SEARCHED

Minimum Documentation Searched 4

| Classification System | Classification Symbols |
|--|---|
| U.S. | 430/172.3, 430/69.7 530/387, 435/172.3 |
| Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched 4 | |

III. DOCUMENTS CONSIDERED TO BE RELEVANT 14

| Category 5 | Citation of Document, 10 with indication, where appropriate, of the relevant passages 11 | Relevant to Claim No. 12 |
|------------|--|--------------------------|
| X Y | US, A, 4,771,002 (Gelvin) 13 September 1988, see column 18, lines 46 to 59, and column 20, lines 48 to 61. | 1-5 6-24 |
| Y | Nucleic Acids Research, Volume 12, Number 9, issued September 1984, Boss et. al. "Assembly of functional antibodies from immunoglobulin heavy and light chains synthesized in <u>E. coli</u> ", pages 3791 to 3806, see entire document. | 6-24 |
| Y | Nature, Volume 312, issued 13 December 1984, Boulianee et. al., "Production of functional chimeric mouse/human antibody", pages 643 to 646, see entire document. | 6-24 |

* Special categories of cited documents: 15

- "A" document defining the general state of the art which is not considered to be of particular relevance.
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search 16

13 November 1990

Date of Mailing of this International Search Report 17

25 JAN 1991

International Searching Authority 18

ISA/US

Signature of Authorized Officer 19

John Ullm
John Ullm

FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

X
Y

Proc. Natl. Acad. Sci. Volume 86, issued July 1989, (U.S.A.), Brenne et. al. "Analysis of mammalian cell genetic regulation in situ by using retrovirus derived portable exons" carrying the Escherichia coli lac Z gene" pages 5517 to 5521, see entire document.

1-5
6-24

V. OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE¹

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. Claim numbers _____, because they relate to subject matter² not required to be searched by this Authority, namely:

2. Claim numbers _____, because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out¹, specifically:

3. Claim numbers _____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING²

This International Searching Authority found multiple inventions in this International application as follows:

1. As all required additional search fees were timely paid by the applicant, this International search report covers all searchable claims of the International application.

2. As only some of the required additional search fees were timely paid by the applicant, this International search report covers only those claims of the International application for which fees were paid, specifically claims:

3. No required additional search fees were timely paid by the applicant. Consequently, this International search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

- The additional search fees were accompanied by applicant's protest.
- No protest accompanied the payment of additional search fees.